

Fi:

```

XX 12-NOV-1993; 93JP-0355504.
XX (MATS/) MATSUBARA K.
XX (OKUB/) OKUBO K.
XX Matsubara K, Okubo K;
XX WPI; 1995-206931/27.
XX
XX Identifying gene signatures in 3'-directed human cDNA library - e.g.
XX for diagnosis of abnormal cell function, by preparing cDNA that
XX reflects relative abundance of corresp. mRNA in specific human
XX tissues
XX
XX Claim 1; Page 1811; 2245pp; Japanese.
XX
XX A single-stranded DNA (or its complementary strand or the corresp.
XX double-stranded DNA) which comprises one of the 7837 "GS" sequences
XX given in AAT19001-T26837 and which is able to hybridize to part of
XX human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
XX sequences were obtained from 3'-directed cDNA libraries prepared
XX from various human tissues; synthesis of cDNA was initiated from the
XX 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
XX all the 3'-oriented cDNAs hybridize with specific mRNA species, almost
XX is constructed so as to reflect accurately the relative abundance of
XX different mRNAs in the particular tissue from which it was derived.
XX The appearance frequency of a given GS in a cDNA library can be
XX determined (esp. using primers and probes derived from the GS
XX sequences) as a means of diagnosing abnormal cell function or for
XX recognising different cell types.
XX
XX Sequence 318 BP; 80 A; 48 C; 66 G; 115 T; 9 other;
XX
XX Query Match 6.8%; Score 264.2; DB 16; Length 318;
XX Best Local Similarity 94.2%; Pred. NO. 2.7e-50;
XX Matches 291; Conservative 0; Mismatches 16; Indels 2; Gaps 2;
XX
QY 3535 gatcttaataccttaatttcattctgtcgaataatgtaactgttttttagaattattacaat 3594
DB 1 gatcttaataccttaatttcattctgtcgaataatgtaactgttttttagaattattacaat 59
XX
QY 3595 atcaatgtgaatatcttgaatctctgttacaatactctgcaactgtattaaacatgtaaatta 3654
DB 60 atcaatgtgaatatcttgaatctctgttacaatactctgcaactgtattaaacatgtaaatta 119
XX
QY 3655 attgttctgtattagccaatctcaccaccacccaatggggagggtatacatgttttgaaaga 3714
DB 20 attgttctgtattagccaatctcaccaccacccaatggggagggtatacatgttttgaaaga 179
XX
QY 3715 c-gtgaactcggtaattgattgttctgtattgttgaactcaatagaagtggtttggaag 3773
DB 180 ctgtgaactcagtaattgattgttctgtattgttgaactcaatagaagtggtttggaag 239
XX
QY 3774 gaagcatggtgtgtgacagtgctgtctcttttggccagctctgtatgtattgttaa 3833
DB 240 gnaagcatggtgtgtgacagtgctgtctcttttggccagctctgtatgtattgttaa 299
XX
QY 3834 gaccatgtt 3842
DB 300 gccntgttt 308
XX
XX RESULT 9
XX AAI10603/c
XX ID AAI10603 standard; DNA; 383 BP.
XX AC AAI10603;
XX DT 12-OCT-2001 (first entry)
XX

```

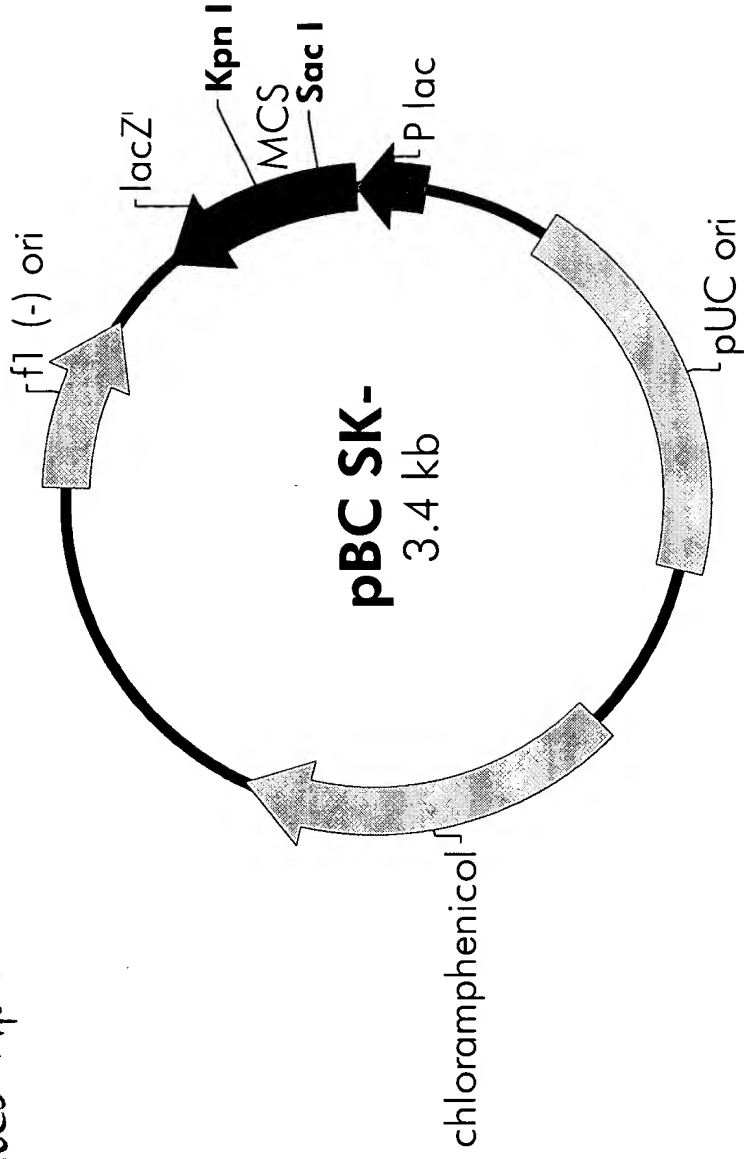
```

DE XX Probe #536 for gene expression analysis in human cervical cell sample.
KW KW Probe: human; microarray; gene expression; cervical epithelial cell;
KW KW cervical cancer; ss.
XX OS Homo sapiens.
XX PN WO200157278-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00670.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -
XX Claim 25; SEQ ID NO 536; 487pp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
XX (SENPs). The present sequence is one such probe. The SENPs are derived
XX from human HeLa cells. The SENPs can be used to produce a single exon
XX microarray, which can be used for measuring human gene expression in a
XX sample derived from human cervical epithelial cells. By measuring gene
XX expression, the probes are therefore useful in grading and/or staging
XX of diseases of the cervix, notably cervical cancer.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 383 BP; 125 A; 71 C; 58 G; 123 T; 6 other;
XX
XX Query Match 5.6%; Score 216.6; DB 22; Length 383;
XX Best Local Similarity 95.8%; Pred. NO. 1.6e-39;
XX Matches 254; Conservative 0; Mismatches 6; Indels 5; Gaps 3;
XX
QY 737 tgacgatctctccagagaagcataaagtgaaactagagtgtctgagaaaaacaccttg 796
DB 383 TGACGATCTCTCCAGAGAGAGCAATAAAGTGAAGTAACTAGAGTCTGAGAAAAACCTTGG 324
XX
QY 797 atccagcttttgatgagacatttacattctatgggtacccctacaccccaaatcc---aag 853
DB 323 ATCCAGCTTTTGATGAGACCTTTTACATTCTATGGGTACCTACACCCAAATCCNNAGA 264
XX
QY 854 aattggccttgcaactcacaat-tttgagttttgacaggttttcaaga-gatgatatcat 911
DB 263 ATTGGNCCCTTGACATTCACAATNTTTGAGTTTGTACAGGTTTTCAGANGATGATATCAT 204
XX
QY 912 tgggggaagtttcaattctctctcgcggaattgaattatctgaagaaaaattttaaataa 971
DB 203 TGGGGAAGTTCTAATTCCTCTCTCGGGAATTGAATTATCTGAGGAAAAAATGTTAATGAA 144
XX
QY 972 tagagagatcatcaagagaaatggt 996
DB 143 TAGAGAGATCATCAAGAGAAATGNT 119
XX
XX RESULT 10
XX AAI1857/c

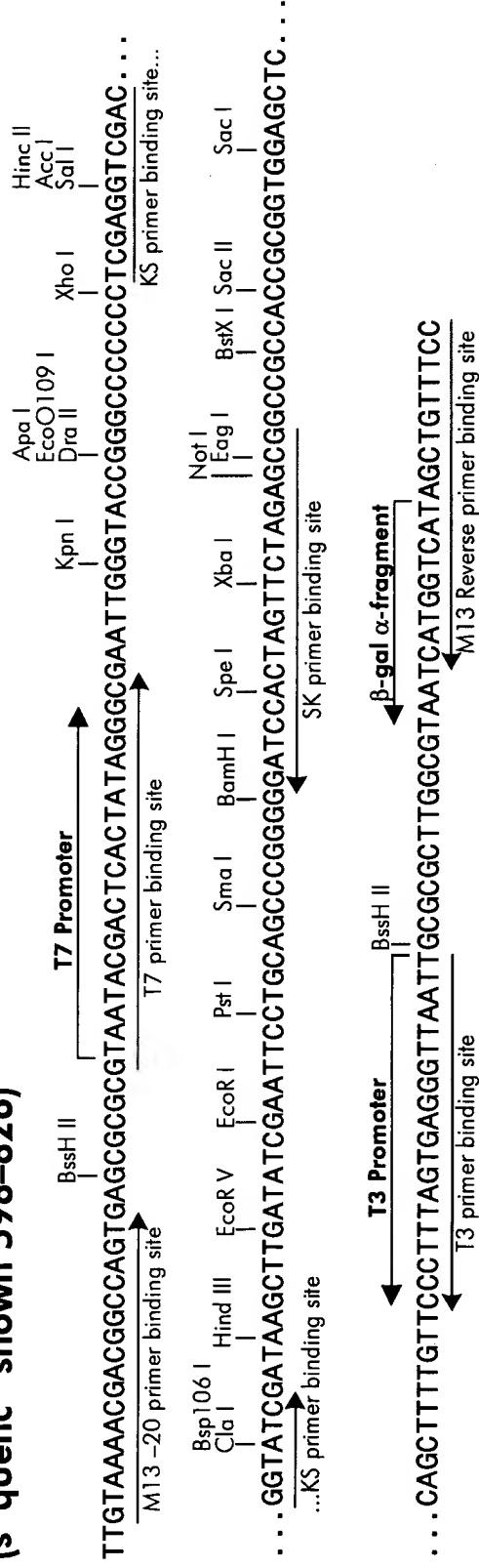
```

p Bluescript+

f1 (-) origin 21-327  
 β-galactosidase α-fragment 460-816  
 multiple cloning site 653-760  
 lac promoter 817-938  
 pUC origin 1158-1825  
 chloramphenicol resistance ORF 1976-2833



### pBC SK (+/-) Multiple Cloning Site Region (sequence shown 598-826)



Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
DNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CCAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LENL at:  
[www.bio.llnl.gov/hbrp/image/html](http://www.bio.llnl.gov/hbrp/image/html)  
Insert Length: 1981 Std Error: 0.00  
Seq primer: -400P from Gibco  
High quality sequence stop: 444.

FEATURES  
source

```

Location/Qualifiers
1. .499
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1947675"
/clone_lib="NCI_CGAP_GC4"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="vector: pT73b-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."
95 c 60 g 173 t
171 a

```

BASE COUNT  
ORIGIN

Query Match 12.2%; Score 474.4; DB 10; Length 499;

[illegible]

RESULT 12  
AA081755

LOCUS 518 bp mRNA EST 21-OCT-1996  
 ZN22D07.r1 Stratagene neuroepithelium NT2RAM1 937234 Homo sapiens  
 CDNA clone IMAGE:548173 5' similar to SW:SYT4\_MOUSE P40749  
 SYNAPTOTAGMIN IV. ; mRNA sequence.  
 AA081755  
 AA081755.1 GI:1623814  
 EST.  
 human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 518)  
 Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,  
 Chissoe, S., Dietrich, N., Dubucq, T., Favello, A., Gish, W., Hawkins  
 , M., Hultman, M., Kucaba, T., Lacy, M., Le, N., Le, N., Mardis, E., Moore  
 , B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,  
 Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E.,  
 Woodward, K., Wohldmann, P., Waterston, R., Wilson, R. and Marra, M.  
 'Generation and analysis of 280,000 human expressed sequence tags  
 Genome Res. 6 (9), 807-828 (1996)  
 9704478  
 Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available royalty-free through LNL ; contact the  
 IMAGE Consortium (info@image.lnl.gov) for further information.  
 Seq primer: -28M13 rev2 from Amersham  
 High quality sequence stop: 309.

[illegible]

```

1. 518
/organism="Homo sapiens"
/db_xref="GDB:3926689"
/db_xref="taxon:9606"
/clone="IMAGE:548173"
/clone_lib="Stratagene neuroepithelium NT2RAMI 937234"
/dev_stage="Ntera-2/RAM+myc neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dr. NT2 (Ntera-2/cl.D1) precursor cells induced with Retinoic Acid for 1 week, followed by 3 weeks in mitotic inhibitor (Replate #2). Average insert size: 1.1 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'
160 a 118 c 112 g 126 t 2 others

```

BASE COUNT  
ORIGIN

Query Match	11.9%;	Score 463.8;	DB 10;	Length 518;
Best Local Similarity	98.5%;	Pred. No. 4.3e-81;		
Matches 509;	Conservative 0;	Mismatches 4;	Indels 4;	Gaps 4;
2Y	68	aaagcgcgcgttgatgtccagcgcaagtgaaatccagccagggcaggttttcccttcacgacc	127	
Db				
	1	AAAGCGCGCGCTTGAGATCCAGCGCAAGTGAATCCAGCCAGGCAGTTTTCCTTCACGACC	60	
2Y	128	tcggcagagaacacgcgcagtgataaaaaatggctccgatcacaccacgcgcggggaagatttgatg	187	
Db				
	61	TCGGACAGAACACGCAGTAAAAAATGGCTCCGATCACACACCGCGGGGAAGAAATTGATG	120	
2Y	188	aaatcccccacagtgfggggattcttcagtgcatttggcgtggtcttcacagttctctctat	247	
Db				
	121	AAATCCCCACAGTGGT-GGGATCTTCAGTGCATTGGCTGGTCTTCACAGTCTCTCTCT	179	
2Y	248	tfgcatggatctgcttcagagaaaaatcatcaagctcaacaagactctccatacaagt	307	
Db				
	180	TTGATGTGATCTGCTGTCTCAGAGAAANTCATCCAGCTTAACAAGACTCCCTCCATCAAGT	239	
2Y	308	tftgcatgtgctttaaggggagtgtgatattaccctgaaacaccttaaatgcaaaaaagaagt	367	
Db				

```

Db 240 TTGTGATGCTTAAAGGAGTTGATATTACCTGAAACCTAAATAGCAAAAAGAGT 299
Qy 368 ttggagcagatgataaaatgaagtaagataaagccagctgtgccaagaattcattgc 427
Db 300 TTGGAGCAGATGATATAAATGAAGTAAAGATAAGCCAGCTGTG-CAAGAGATTCATTGC 358
Qy 428 atctgagattgaaagagagatctcaatggaatttcccaaaaaccacacacacacac 487
Db 359 ATCTGATCTTTGAAAGAGAGATCTCAATGGCAATTTTCCCAAAACCAACNTCAACCTG 418
Qy 488 gacgtccctctgagctgagaaatgaacccggaagctcttttttagaaggggaaagag 547
Db 419 GCAGTCTTCTGATCTGAGAAATGCAACCCGAGACTCTTTTAGAAGGGGAAAGAGT 478
Qy 548 cagtttccctcagagagtttaagtcagacactccct 584
Db 479 CAGTTT-CCCTGAGAGTTT-AGTCCGACACTTNCCT 513

RESULT 13
AV606332 565 bp mRNA EST 30-AUG-2000
DEFINITION AV606332 Bos taurus kidney fetus Bos taurus cDNA clone EIK1034B06
5', mRNA sequence.
ACCESSION AV606332
VERSION AV606332.1 GI:9736705
KEYWORDS EST.
SOURCE COW.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 565)
AUTHORS Sugimoto,Y., Hirotsune,S., Takasuga,A., Itoh,R., Jitohzono,A. and
Suzuki,H.
JOURNAL bovine cDNA sequencing
COMMENT Unpublished (2000)
Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugi@ccoc.ocn.ne.jp
Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.
FEATURES
Source
1..565
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="EIK1034B06"
/clone_lib="Bos taurus kidney fetus"
/tissue_type="kidney"
/dev_stage="fetus"
/lab_host="DH10B"
/notes="vector: pZLI; Site_1: SalI; Site_2: NotI; Poly A
was deleted from a NotI site"
BASE COUNT 162 a 110 c 144 g 148 t 1 others
ORIGIN
1
11.88; Score 458; DB 10; Length 565;
Best Local Similarity 90.3%; Pred. No. 5.9e-80;
Matches 501; Conservative 0; Mismatches 51; Indels 3; Gaps 1;

Qy 915 ggaagtcttaattctctcgcgaattgaattatctgaaggaataatgttaataagtag 974
Db 11 GGAAGTCTGATTCCTCGCAGGAATGAATTAACATAAGGGAATGTAAATGAACAG 70
Qy 975 agagatcatcaagaagaatgttagaagcttccagagcggggtgagttactgactctct 1034
Db 71 AGAGATTACCAAGAGAATGTAGGAATCTTCAGGACGGGGTATGATCTCTCT 130

```

```

Qy 1035 ctgctatcagtcaccacacacacactctaaactgtggtgtcttaaaagctcgacatctgcc 1094
Db 131 CTGCTACCACTGCACACAAATACTCTTACTGTGGTGTGTTTAAAGCTCGGCACCTGCC 190
Qy 1095 taaatctgattgtccggagactttcagatccctatgtcaaaagtgaacctgtaccatgccaa 1154
Db 191 TAAATCTGATGTCTGTGAGCTTTTCAGATCCCTATGTCAAAAGTGAACCTGTACCATGCCAA 250
Qy 1155 aaagagaatctcacaagaagaagactcatgtgaagaatgcaccccccaatgcagtttcaa 1214
Db 251 AAAGAGAATCTCTAAAGAAAGAGACACATGTGAAGAAATGCACGCCCAATGCAGTTTCAA 310
Qy 1215 tgagctgttctgtcttgatattccttgagggccttgagataataagtttgaatttt 1274
Db 311 TGAACCTCTTGTCTTGACATTCCTTGTGAGGCTCTAGAAAGAGATGAAGTTTGAATTTCT 370
Qy 1275 ggtttggattctgaaaggggggtcccgaaatgaggttaactcggcagcttagctctgggtgc 1334
Db 371 GGTTTGGATTCCTGAAGGGGATCCCGAAATGAGTGAATGGGGGGTGTGCTCTGGGAGC 430
Qy 1335 agcagcagaaggaactggtggagagcactggaagaagatctgtgactacccaggagaca 1394
Db 431 AGCAGCAGAAGGAGCTGCTGGAGAGCAGCTGGAAGAGAGATCTGTGACTATCCAGAGGCA 490
Qy 1395 aattgccaaagtgcagctgctctgtgaggttagcttagcctcctagccg---tgagttggaact 1451
Db 491 AATCGCAAGTGGCAGCAGCTCTGTGATGTTAGCACCCCTTANCCGTGATGAATTTGGAAC 550
Qy 1452 taaaggttttacta 1466
Db 551 TAACGATTTTACTA 565

RESULT 14
AW914163 712 bp mRNA EST 25-MAY-2000
LOCUS EST345467 Normalized rat brain, Bento Soares Rattus sp. cDNA clone
DEFINITION RGIAB26 5' end, mRNA sequence.
ACCESSION AW914163
VERSION AW914163.1 GI:8079837
KEYWORDS EST.
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 712)
AUTHORS Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
Kerlavage,A.R. and Adams,M.D.
Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat
Gene Index
JOURNAL Unpublished (1998)
COMMENT Other_ESTS: EST345466
Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
This clone is available through the ATCC, contact the ATCC
tel#703-365-2700 for further information
Seq primer: M13 Reverse.
FEATURES
Source
1..712
/organism="Rattus sp."
/db_xref="taxon:10118"
/clone="RGIAB26"
/clone_lib="Normalized rat brain, Bento Soares"
/notes="Organ: brain; Vector: pT73Pac; Site_1: EcoRI;
Site_2: NCI"
BASE COUNT 185 a 145 c 178 g 204 t
ORIGIN

```